DOCUMENT RESUME

ED 109 179 TM 004 627

AUTHOR Subkoviak, Michael J.; Roecks, Alan L.

TITLE Nonmetric Multidimensional Scaling: An Evaluation of Three Data Collection Methods.

PUB DATE [Apr 75]

NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D. C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE

DESCRIPTORS *Data Collection; *Distance; *Evaluation Methods; Higher Education; *Multidimensional Scaling; *Nonparametric Statistics; Rating Scales; Statistical Analysis

ABSTRACT

The present study examined three different methods of data collection in which subjects judged proximity between object pairs. One method required subjects to partition objects into homogeneous subsets; the second entailed rating object pairs on a similarity-dissimilarity continuum; and the third involved comparing inter-object proximities to a fixed standard. The three types of proximities were analyzed by the nonmetric multidimensional scaling procedure, and subsequent multidimensional representations were compared for accuracy to a criterion or "true" multidimensional configuration of the same objects generated by the same subjects. Considerable differences in accuracy were found among the methods. (Author)

ED109179

Nonmetric Multidimensional Scaling: An Evaluation of Three Data Collection Methods

Michael J. Subkoviak

Alan L. Roecks

The University of Wisconsin

Abstract

The present study examined three different methods of data collection in which subjects judged proximity between object pairs. One method required subjects to partition objects into homogeneous subsets; the second entailed rating object pairs on a similarity-dissimilarity continuum; and the third involved comparing inter-object proximities to a fixed standard. The three types of proximities were analyzed by the nonmetric multidimensional scaling procedure, and subsequent multidimensional representations were compared for accuracy to a criterion or "true" multidimensional configuration of the same objects generated by the same subjects. Considerable differences in accuracy were found among the methods.

TM 004 627

Nonmetric Multidimensional Scaling: An Evaluation of Three Data Collection Methods

Michael J. Subkoviak

Alan L. Roecks

The University of Wisconsin

Background

Given a measure of proximity (similarity or dissimilarity) between each pair of n objects $o_1, o_2, \ldots, o_n$, nonmetric multidimensionsl scaling (Kruskal, 1964a, 1964b; Shepard, 1962a, 1962b) represents the objects as points in t-dimensional space so that distances between pairs of points are monotonically related to corresponding input proximities. The proximities need only be distance-like (Shepard, 1972, p. 24) and measured on at least an ordinal scale, thus methods of generating such data are quite varied.

Due to the definitional generality of "proximity", an exhaustive listing of data suitable for nonmetric multidimensional scaling is all but impossible. However Coombs (1964), Torgerson (1958) and Wish (1972) have compiled taxonomies of commonly used proximity measures and methods for collecting them.

Since behavioral sciences focus on the activity of animate entities, proximities are commonly generated by having subjects employ their cognitive, affect or psychomotor facilities in judging inter-object similarities and dissimilarities. The present study deals with three such judgmental methods.

Non-judgmental techniques constitute a second source of proximity measures.

This class includes methods in which objectively measured properties of objects are used to derive a proximity index, e.g. two objects can be measured on height, weight and other relevant attributes and the correlation or sum of squared, standardized differences across these variables taken as the inter-object proximity. Non-judgmental indices are also employed within the behavioral sciences and are particularly common to outside disciplines. The present study does not involve non-judgmental proximities.

Inter-object proximities were judged by subjects in the present study in the following ways: (1) proximity grouping, (2) proximity rating and (3) proximity comparison. Each of these methods is described in detail below.

Proximity grouping. The instructions in proximity grouping tasks ask subjects to sort n objects into mutually exclusive and exhaustive groups so that objects in the same group are more similar to each other than to objects in other groups. The proportion (or frequency) of times objects o_i and o_j are sorted into the same group is often used as a measure of proximity. This type of index has been employed in multidimensional analysis of personality traits (Rosenberg, Nelson & Vivekananthan, 1968), nations (Wish, Deutsch & Biner, 1970) and university faculty (Subkoviak & Levin, 1974).

The judgment required in proximity grouping is quite simple, and thus the method is particularly appropriate for use with complex object properties or unsophisticated subjects. Another advantage of the procedure is that subjects can respond to a large number of objects in a relatively short span of time.

A distinct disadvantage of this approach is that a single sorting provides no information about proximity differences between objects within the same group or about proximity differences between groups. These shortcomings can

be remedied, but at the expense of time and simplicity. For example, after the initial grouping, subjects can be asked to judge proximities between different groups or between objects within the same group.

Proximity rating. Subjects are presented with all possible object pairs (o_i,o_j) and are asked to rate the proximity of each pair on a scale, e.g. "SIMILAR : 0 1 2 3 4 5 6 7 8 9 : DISSIMILAR." The average scale value of pair (o_i,o_j) across all ratings is often used as a proximity measure. This type of index has been employed in multidimensional analysis of geometric figures (Attneave, 1950), attitudes (Messick, 1954) and interpersonal relations (Wish, Kalplan & Deutsch, 1974).

A number of variations in mode of object presentation and type of rating scale are possible (Torgerson, 1958; Wish, 1972); but the judgmental task remains basically one of judging the absolute proximity of each object pair, as opposed to judging the proximities of various pairs relative to one another (see the proximity comparison method discussed below).

Complete information is obviously obtained about all n(n-1)/2 inter-object proximities, at the expense of time and subject fatigue as n becomes large. More important, the process of averaging ratings across subjects assumes that the scale has the same meaning for all subjects. This is typically a somewhat over-optimistic supposition in the light of subjects' variable response tendencies, e.g. attraction to or avoidance of the extreme ends of a rating scale.

Proximity comparison. Subjects are asked to judge the degree of proximity between object pairs (o_i,o_j) relative to a standard. For example, subjects might report the percent of similarity between (o_i,o_j) compared to the similarity between a standard object-pair, and the average percent for (o_i,o_j) across

all subjects taken as a proximity measure. This approach has been used with much success in judging the geographic proximity of various world cities (Lundberg & Ekman, 1973), and the method generalizes easily to other stimuli.

Other variations of the comparative judgment approach, which are not considered in the present study, are possible. For example, subjects might be asked to provide a complete or weak (ties allowed) ranking of pairs (o_i,o_j) from most similar pair to least similar pair, and the average rank of (o_i,o_j) across subjects taken as a proximity measure (Rapoport & Fillenbaum, 1972). However, as the number n and homogeniety of objects increase, this task becomes quite difficult. Torgerson (1958) and Coombs (1964) also describe indirect procedures for transforming object comparisons to proximity data.

Like rating techniques, comparison methods produce complete information about inter-object proximities for a greater investment of time. In addition, such procedures may produce more valid and consistent data since comparisons among actual objects tend to be better-defined than ratings on a somewhat ambiguous scale.

Method

Subjects. A total of 600 undergraduate and graduate students enrolled in communication arts and educational psychology courses during the Spring Semester of 1974 at the University of Wisconsin took part in the study.

Materials. The present study compared the accuracy of the aforementioned sorting, rating and comparing methods in judging distances (geographic proximities) between U.S. cities (objects). Two sets of ten cities were considered: (1) Set 1 ≡ {Philadelphia, Baltimore, Detroit, Atlanta, Chicago, New

Orleans, Denver, Phoenix, Seattle, Los Angeles} and (2) Set 2 ≡ {Detroit, Cincinnati, Atlanta, Minneapolis, St. Louis, Kansas City, New Orleans, Denver, Phoenix, Houston}. The 10(10-1)/2 = 45 inter-city distances of Set 1 are heterogeneous (standard deviation 660 miles) while those of Set 2 are homogeneous (standard deviation 365 miles). Thus inter-city distances of Set 1 are generally easier to differentiate than those of Set 2, and consequently judged proximities for Set 1 tend to be more consistent with an individual's cognitive map than for Set 2.

Questionnaires were constructed for each of the three data collection methods, one involving the cities of Set 1 (shown below) and the second involving those of Set 2 which were listed at the top of the questionnaire.

The directions for the proximity grouping procedure were as follows.

> Sort the cities into separate groups in the blank space below, so that cities in the same group have small distances between them and are near one another. Please sort each city into one and only one group. Draw a circle around each separate group. Use as few or as many groups as you think are necessary; each group may contain as few or as many cities as seem appropriate.

The instructions for the proximity rating method were as follows.

> In Item 1 below rate the distance between Detroit and Los Angeles on the scale 0,1,2,3,4,5,6,7,8,9. Small numbers 0-4 indicate small distances, and the smaller the number the smaller the distance between the two cities. Large numbers 5-9 indicate large distances, and the larger the number, the larger the distance between the two cities.
>
> Please circle one and only one number 0,1,2,3,4,5,6,7,8,9 in each of the following items.

The 45 possible pairs of cities were then listed in random order, the order within each pair also being randomized.

The directions for the proximity comparison technique were as follows.

> Imagine that the distance between New York City and San Francisco equals 100 units. Now compare the distances between the cities above to the distance between New York City and San Francisco. For example, the distance between Detroit and Los Angeles is what percent of the distance between New York City and San Francisco? Record your answer in Item 1 below.
>
> Complete all the other items in the same way. Compare the distance between the given cities to the distance between New York City and San Francisco, and then record your answer as a percent.

The 45 pairs of cities were again listed in the same randomized order as for the proximity rating method.

After all proximity grouping, rating or comparing judgments were made, subjects were instructed to break the seal on the last page of the questionnaire and to make a copy of their cognitive map on a completely blank outline representation of the Continental United States (no man-made or natural features of any kind were depicted). The inter-city distances on this "true" map were used to determine the accuracy of the grouping, rating or comparing judgments of those same distances. The actual instructions were as follows.

> Use your mental picture or image of the U.S. to locate the following cities on the map provided below.

1.	Los Angeles	6.	Chicago
2.	Phoenix	7.	Denver
3.	Baltimore	8.	New Orleans
4.	Detroit	9.	Atlanta
5.	Seattle	10.	Philadelphia

> Use a dot (•) to indicate the location of each city on the map below. Then write the number of each city (1 thru 10 above) over its dot ($\overset{1}{\bullet}$). Please be sure to place all 10 cities on the map.

Procedure. The three data collection methods were completely crossed with the two sets of cities for a total of six conditions, and 100 subjects were randomly assigned to each condition. Subjects were tested in groups of about

20-150. The six forms of questionnaire were arranged in cyclical order and distributed one at a time, randomizing the assignment of subjects to conditions (Underwood, 1966, p. 115). Subjects then completed the sorting, rating or comparing judgments and the cognitive map. The amount of time taken to complete the questionnaires was recorded.

Analysis

A numerical judgment S_{ij} of the proximity between the 45 possible pairs of cities was obtained for each condition, small numbers indicating that a pair was geographically close and large numbers meaning the opposite. In the sorting task a pair was coded 0 if a subject placed those two cities in the same group or 1 if they were placed in different groups; in the rating task pairs were scored 0 thru 9; and in the comparing task pairs were generally scored 0 thru 100 percent (a few pairs were judged greater than 100 percent of the standard by a small number of subjects). Scores S_{ij} for pair (o_i, o_j) were then averaged across the 100 subjects in each condition to obtain a group proximity measure $\overline{S}_{ij} = \Sigma S_{ij}/100$ for the pair.

The proximity measures $\overline{S}_{ij}$ were next input into a nonmetric multidimensional scaling program MINISSA-I (Lingoes, 1973), and a two-dimensional representation of the ten cities (defined by numerical coordinates) was obtained for each condition. The 45 Euclidean distances $d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2}$ were computed between all pairs of cities in the representation, where (x_{i1}, x_{i2}) are the two-dimensional coordinates locating city o_i in the configuration for the 100 subjects. The purpose of the study was to compare these MINISSA-I distances for accuracy to the 45 corresponding true distances between cities on the cognitive maps of the same 100 subjects.

Accordingly, the true coordinates (X_{i1},X_{i2}) locating each city o_i on each of the 100 cognitive maps were obtained using an electronic digitizer that determines coordinates in units of 1/200$^{\text{ths}}$ of an inch. True Euclidean distances $D_{ij} = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2}$ were obtained for each subject and averaged to provide a group measure $\bar{D}_{ij} = \Sigma D_{ij}/100$ of the true distance between pair (o_i,o_j).

For purposes of comparison, the MINISSA-I distances d_{ij} were transformed to the same units of measurement (1/200$^{\text{ths}}$ of an inch) as true distances $\bar{D}_{ij}$. The new MINISSA-I distances were given by $d'_{ij} = a \cdot d_{ij}$ where $a = \Sigma d_{ij}\bar{D}_{ij}/\Sigma d^2_{ij}$ is chosen to minimize $\Sigma(\bar{D}_{ij} - d'_{ij})^2$. This is an admissible linear transformation of ratio scale distances (Lord & Novick, 1968, p. 21) and corresponds to a uniform shrinking of the MINISSA-I configuration to make it comparable to the cognitive map.

Finally, the typical percent of discrepancy between true $\bar{D}_{ij}$ and MINISSA-I d'_{ij} across all 45 distances was computed as the measure of correspondence between the actual cognitive maps and the sorting, rating or comparing judgments of the 100 subjects in a given condition (Kruskal, 1964a, p. 15).

$$\text{Percent Error} = \sqrt{\frac{\sum_{1}^{45} [\bar{D}_{ij} - d'_{ij}]^2}{\sum_{1}^{45} [\frac{\bar{D}_{ij} + d'_{ij}}{2}]^2}}$$

For example, as shown in Table 1 for sortings of Set 1 cities by N = 100 subjects, d'_{ij} typically differ from true $\bar{D}_{ij}$ by 32 percent.

Finally, to determine if results are consistent for various size groups, the 100 subjects in each condition were randomly partitioned into five groups of N = 20 and independently into ten groups of N = 10 subjects; and the analysis described above was repeated for each of the smaller groups.

Results and Discussion

Table 1 shows the percent of discrepancy between true $\overline{D}_{ij}$ and MINISSA-I d'_{ij} for various size groups N = 100, 20, 10 under the six method x stimuli conditions; for group size N = 20 and N = 10, the median percent error across five and ten subgroups respectively is reported. A number of conclusions can be drawn from these results.

Insert Table 1 about here

First, for the heterogeneous stimuli of Set 1 and for a given group size N, sorting distances d'_{ij} contained about four times more error than either rating or comparing distances. Thus the tendency of subjects to sort Set 1 cities into the same clusters and the resulting lack of proximity information about cities within a cluster introduced considerable error into the nonmetric multidimensional scaling configuration. As one would expect, this condition became more severe with decreasing group size and loss of intra-cluster information due to the increased likelihood of complete agreement among subjects' sortings. Increased agreement among the ratings and comparisons of smaller groups and the slight reduction in discriminations among like proximities had a similar effect on the configurational accuracy of these tasks.

Second, for the homogeneous stimuli of Set 2, the sorting, rating and comparing methods were equally accurate for applied purposes, except for small samples N ≤ 10; in fact, the sorting method produced noticeably more accurate results for Set 2 than Set 1 due to subjects' greater variability in defining Set 2 clusters. The former result has interesting practical implications. In the behavioral sciences, a proximity S_{ij} is normally perceived

variably across subjects, as for Set 2 stimuli. Under such conditions, the sorting procedure tends to produce results as accurate as rating or comparing, assuming $N > 10$. Furthermore, sorting requires considerably less time than rating or comparing. For only $n = 10$ objects in the present study, the average sorting time was 7 1/2 minutes as opposed to 11 and 11 1/2 minutes for rating and comparing. Since sorting time is a function of n while rating and comparing time are a function of $n(n-1)/2$, this difference grows rapidly as n increases.

References

Attneave, F. Dimensions of similarity. American Psychologist, 1950, 63, 516-556.

Coombs, C. H. Theory of Data. New York: Wiley, 1964.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964, 29, 1-27. (a)

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964, 29, 115-129. (b)

Lingoes, J. C. The Guttman-Lingoes Nonmetric Program Series. Ann Arbor, Mich.: Mathesis Press, 1973.

Lord, F. M. & Novick, M. R. Statistical Theories of Mental Test Scores. Reading, Mass.: Addison-Wesley, 1968.

Lundberg, U. & Ekman, G. Subjective geographic distance: A multidimensional comparison. Psychometrika, 1973, 38, 113-122.

Messick, S. J. The Perception of Attitude Relationships: A Multidimensional Scaling Approach to the Structuring of Social Attitudes. Doctoral Dissertation, Princeton University, 1954.

Rapoport, A. & Fillenbaum, S. An experimental study of semantic structures. In A. K. Romney, R. N. Shepard and S. B. Nerlove (Eds.), Multidimensional Scaling. Vol. II. New York: Seminar Press, 1972, pp. 93-131.

Rosenberg, S., Nelson, C. & Vivekananthan, P. S. A multidimensional approach to the structure of personality impressions. Journal of Personality and Social Psychology, 1968, 9, 283-294.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. Psychometrika, 1962, 27, 125-140. (a)

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. II. Psychometrika, 1962, 27, 219-246. (b)

Shepard, R. N. In R. N, Shepard, A. K. Romney and S. B. Nerlove (Eds.), Multidimensional Scaling. Vol. I. New York: Seminar Press, 1972. Pp. 21-47.

Subkoviak, M. J. & Levin, J. R. Determining the characteristics of the ideal professor: An alternative approach. Journal of Educational Measurement, 1974, in press.

Torgerson, W. S. Theory and Methods of Scaling. New York: Wiley, 1958.

Underwood, B. J. Experimental Psychology. New York: Appleton-Century-Crofts, 1966.

Wish, M. Notes on the variety, appropriateness and choice of proximity measures. Paper presented at the Bell-Penn Workshop on Multidimensional Scaling, Philadelphia, June, 1972.

Wish, M., Deutsch, M. & Biner, L. Differences in conceptual structures of nations: An exploratory study. Journal of Personality and Social Psychology, 1970, 16, 361-373.

Wish, M., Kaplan, S. J. & Deutsch, M. Dimensions of interpersonal relations: Preliminary results. Paper presented at the 81st Annual Convention of the American Psychological Association, Montreal, 1973.

Table 1

Percent Error Between True Cognitive Map and Derived Nonmetric Multidimensional Scaling Configuration

Method	Sorting			Rating			Comparing		
Stimuli	N = 100	N = 20	N = 10	N = 100	N = 20	N = 10	N = 100	N = 20	N = 10
Set 1	.32	.40	.42	.09	.12	.14	.08	.09	.10
Set 2	.14	.21	.29	.17	.16	.17	.15	.13	.15